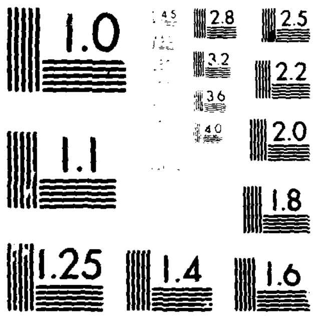
AD-A084 237 ILLINOIS UNIV AT URBANA-CHAMPAIGN DEPT OF PSYCHOLOGY F/G 5/9
A COMPARISON OF SINGLE- AND DUAL-TASK MEASURES TO PREDICT PILOT--ETC(U)
MAY 79 D L DAMOS, G LINTERN F44620-76-C-0009
UNCLASSIFIED ENG-PSY-79-2/AFOSR-79-2 AFOSR -TR-80-0325 NL

1 OF 1
AD
A084237

END
DATE
FILMED
6-80
DTIC



MICROCOPY RESOLUTION TEST CHART
NATIONAL BUREAU OF STANDARDS-1963-A

LEVEL II

University of Illinois at Urbana-Champaign
Department of Psychology

Technical Report
Eng Psy-79-2/AFOSR-79-2

May 1979

A COMPARISON OF SINGLE- AND DUAL-TASK MEASURES TO PREDICT PILOT PERFORMANCE

Diane L. Damos
and
Gavan Lintern

Prepared For :

Air Force Office of Scientific Research
Air Force Systems Command
United States Air Force

unclassified

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE

READ INSTRUCTIONS BEFORE COMPLETING FORM

18 19

1. REPORT NUMBER: [redacted]FØSR-TR-80-0325

2. GOVT ACCESSION NO.: AD-A084 237

3. RECIPIENT'S CATALOG NUMBER:

4. TITLE (and Subtitle): 6 A COMPARISON OF SINGLE- AND DUAL-TASK MEASURES TO PREDICT PILOT PERFORMANCE,

5. TYPE OF REPORT & PERIOD COVERED: 9 Technical Rep[redacted]t,

6. PERFORMING ORG. REPORT NUMBER: Eng-Psy-79-2/AFOSR-79-2

7. AUTHOR(s): 10
Diane L. Damos SUNY at Buffalo
Gavan Lintern Canyon Research Group, Inc.

8. CONTRACT OR GRANT NUMBER(s): F44620-76-C-0009 15

9. PERFORMING ORGANIZATION NAME AND ADDRESS:
University of Illinois
Department of Psychology
Urbana, Illinois 61801

10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS: 61102F, 2313 A4 16 17

11. CONTROLLING OFFICE NAME AND ADDRESS:
U.S. Air Force
Air Force Office of Scientific Research (NL)
Building 410
Bolling AFB, DC 20332

12. REPORT DATE: 11 May [redacted]79

13. NUMBER OF PAGES: 24

14. MONITORING AGENCY NAME & ADDRESS (if different from Controlling Office): 12 29

15. SECURITY CLASS. (of this report): Unclassified

15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report):
Approved for public release; distribution unlimited

14 ENG-PSY-79-2/AFOSR-7

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

18. SUPPLEMENTARY NOTES

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)
Human Timesharing
Dual-Task Performance
Prediction of Flight Performance

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

An experiment comparing the predictive validity of single- versus dual-task measures is reported. Fifty-seven males received two trials on each of two identical one-dimensional compensatory tracking tasks. The subjects then attempted to perform the tasks concurrently for 25 trials. Finally, they performed each task alone for one trial. The subjects then were given a short basic flight course consisting of ground instruction and practice in a GAT-2 simulator. After completing the course, the subjects

DD FORM 1 JAN 73 1473

unclassified

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

4 159

were asked to perform four repetitions of a descent, a descent followed by a stall, and a level turn. Performance was scored by an instructor and an observer. Performance in the simulator then was correlated with performance on each tracking trial. The predictive validity of the early single-task scores decreased with practice while the dual-task validity increased throughout the testing session. However, the predictive validity of the late single-task scores was almost as large as that of the late dual-task scores. Possible explanations for the results are given.

University of Illinois at Urbana-Champaign
Department of Psychology

Technical Report
Eng Psy-79-2/AFOSR-79-2

May 1979

A COMPARISON OF SINGLE- AND DUAL-TASK MEASURES TO PREDICT PILOT PERFORMANCE

Diane L. Damos

and

Gavan Lintern

Prepared for
AIR FORCE OFFICE OF SCIENTIFIC RESEARCH
AIR FORCE SYSTEMS COMMAND
UNITED STATES AIR FORCE

TABLE OF CONTENTS

INTRODUCTION

Multiple-task performance measures occasionally have been used to predict success in flight training. The rationale behind the use of these scores is that they measure timesharing skills, the skills involved in performing two or more tasks concurrently. It has been assumed that timesharing skills contribute substantially to operational effectiveness and that they are different from those skills measured by paper-and-pencil tests and single-task psychomotor tests. Measures of timesharing skill, therefore, could improve the overall predictive validity of a selection battery for advanced stages of flight training.

The first major investigation of multiple-task measures as predictors of success in pilot training was undertaken by the Army Air Force in World War II (Melton, 1947). Performance on the pursuit rotor task with divided attention was used to predict success in subsequent stages of pilot training. The predictive validity of this task was quite low, approximately .20, considerably lower than the predictive validities of several other apparatus tests such as the complex coordination test (.40).

Later attempts to predict pilot performance from multiple-task measures have been more successful. Trankell (1959) has described a timesharing test used by the Scandanavian Airline System to select pilots for their training program. The test consisted of two patterns of circles connected with straight lines. One pattern was for the right hand and one for the left. The candidate held a pencil in each hand and moved the pencils alternately from circle to circle along the lines to the beat of a metronome while solving a number of intellectual problems. A subjective rating of performance in this test had the highest correlation with success in pilot training of any test administered (.42) based on a sample of 363 experienced pilot applicants.

North and Gopher (1976) tested 32 student pilots at the commencement of a private pilot course. Each student performed a one-dimensional compensatory tracking task and a choice reaction time task alone and then together. The system dynamics of the tracking task were adapted to correct for individual differences in timesharing skills. Measures of single-task and multiple-task performance were correlated with instructor's assessments of student capabilities. Only multiple-task measures correlated reliably with these ratings.

Jacobs (1976) examined the effects of three types of flight simulator motion on transfer of basic flight skills to the aircraft. Each of 36 subjects was assigned to one of four groups on the basis of a composite performance score from the Gopher and North (1976) dual-task test. The composite consisted of normalized performance scores on four measures: the acceleration component in adaptive tracking, correct response time, the ratio of the single- to dual-task RMS error, and the ratio of single- to dual-task correct response time. The composite score correlated .40 with trials to criterion in the aircraft ($p < .05$), .45 with errors to criterion ($p < .01$), and .47 with time to criterion ($p < .01$).

Damos (1978) examined 33 students enrolled in a private pilot course. Before these students began training, they simultaneously performed a one-dimentional compensatory tracking task and a choice reaction time task at 1, 2, and 3 bits of stimulus information. Cross-adaptive logic was used to keep performance on the tracking task within narrow error limits, thereby casting the variance associated with individual differences into the reaction time measure.

The Illinois Private Pilot Flight Performance Scale (Povenmire, Alvares, and Damos, 1970) was administered to all students after 10, 20, and 30 hours of flight training. The multiple correlations between scores on this scale after each period of flight training and reaction times at the three levels of information were .59, .63, and .68. Of these three correlations only the last is statistically reliable. However, it should be noted that the correlation increased as a function of interpolated flight training.

Although the studies by Trankell (1959), North and Gopher (1976), Jacobs (1976), and Damos (1978) suggest that performance on a multiple task can predict success in training, these studies have a number of shortcomings. First, neither Trankell nor Damos assessed single-task performance on either of the component tasks. Thus, the predictive correlations could reflect the relation between the component skills and the criterion behavior. Second, although Jacobs' composite did correlate reliably with three measures of aircraft performance, it contained dual- as well as single-task measures. Subsequent analyses indicated that only single-task tracking scores correlated reliably with two of the measures. Third, although North and Gopher found reliable correlations between instructor ratings of student ability and dual-task performance measures, they did not examine actual flight performance. Finally, none of these four studies demonstrated that timesharing skills actually were required to perform the test battery. Since one purpose of multiple-task testing is to assess timesharing skills, it seems necessary to demonstrate that these skills are required by the test combination.

The experiment described in this paper compares the predictive validities of single- and dual-task measures when timesharing skills are involved demonstrably in the performance of the dual task.

IDENTIFICATION OF TIMESHARING SKILLS

A technique developed and more fully described by Damos (1977) was used to isolate timesharing skills. This technique, which is illustrated in Figure 1, divides training into two stages. During Stage 1, which involves single-task training, each component task is practiced until performance has stabilized. In Stage 2, which is predominately dual-task training, single-task performance is reassessed periodically to determine its stability. If dual-task performance improves during Stage 2 while single-task performance remains stable, the improvement may be attributed to the development of timesharing skills.

To demonstrate statistically the development of timesharing skills using this technique, a two-factor (secondary-task load by trials) analysis of variance is applied to the Stage 2 data. Both the main effect of secondary-task load and the interaction of secondary-task load with practice must be statistically reliable. The main effect indicates a dual-task decrement in performance and the interaction in conjunction with stable single-task performance implies that improvement in dual-task performance is the result of improved timesharing skills and not of improved single-task skills.

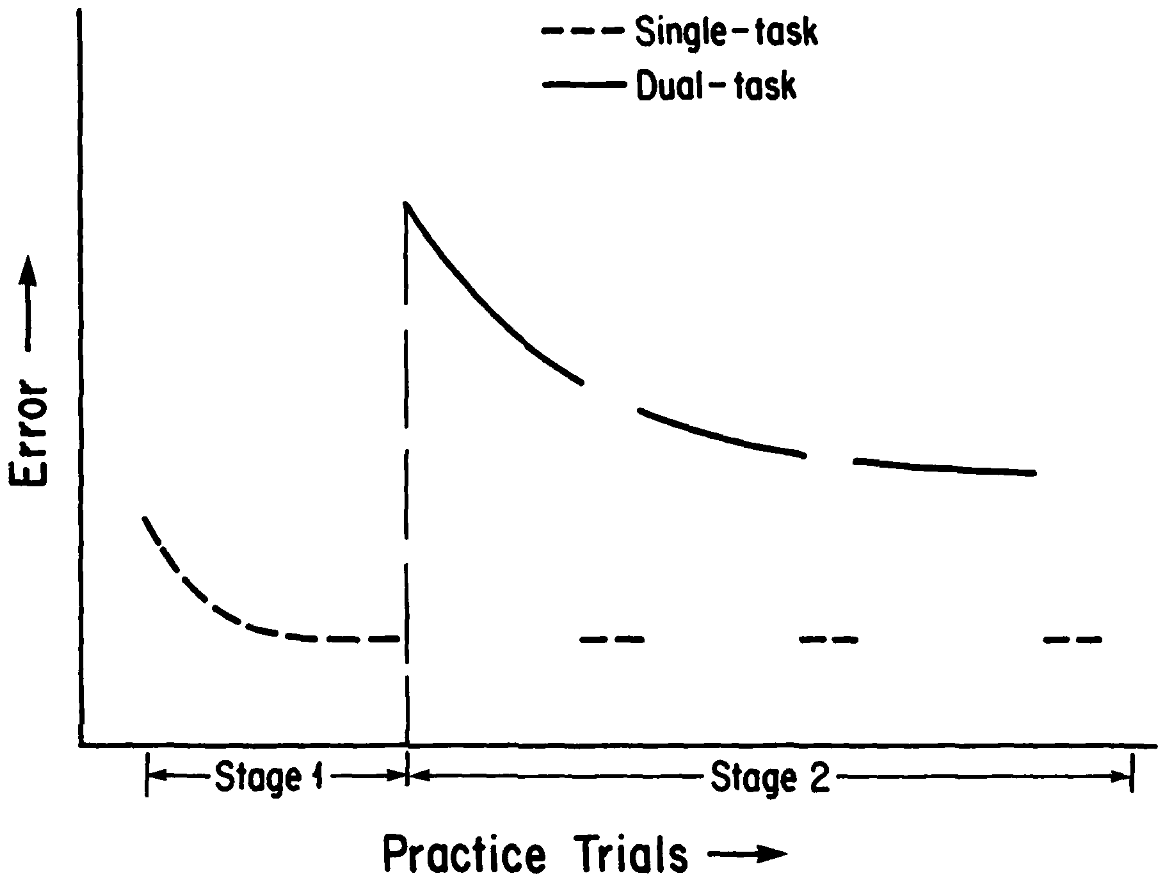


Figure 1. An example of the measurement technique used to identify time-sharing skills. During Stage 1 practice on each component task is continued until performance begins to stabilize. During Stage 2, practice is primarily on the dual-task combination. However, single-task performance is periodically reassessed to determine its stability.

METHOD

Subjects

A total of 66 flight-naive male volunteers between the ages of 18 and 30 were selected to participate in the experiment from those who replied to advertisements in a student newspaper and posters placed in various university buildings. These subjects were recruited to participate in an experiment on the effects on augmented feedback on learning (Lintern, 1978). Because the Lintern study required the subjects to reach criterion on several different maneuvers in a simulator, it was of interest to eliminate individuals who could not reach criterion before they began the experiment. Two of the tasks reported in this paper, dual-task tracking and the Bennett Test of Mechanical Comprehension, were used as selection devices for Lintern's experiment. A cutting score was established for the Bennett Test of Mechanical Comprehension based on published test norms and the results of previous work by Damos (1977). Five subjects had scores below this cutoff and were subsequently eliminated. Four more subjects had scores below a cutoff point on the dual-task tracking, which again was established based on previous work by Damos (1977).

Tasks

Tracking test. Two identical one-dimensional compensatory tracking tasks required the subject to keep a moving circle centered in a horizontal track by making appropriate left-right manipulations of a control stick. One task was controlled by each hand. Under single-task conditions the display was centered on the face of a CRT. Under dual-task conditions the tracking task controlled by the subject's right hand was presented slightly to the right of the center of the display and below the left-hand task, which was displaced slightly to the left of the center of the display. The visual angles subtended by the dual-task display were 4.05° horizontally and 0.70° vertically (.07 by .01 rad). The inputs to the tasks were independent random forcing functions with upper cutoff frequencies of .32 Hz. The control systems had identical mixed first- and second-order dynamics with weightings of .10 and .90 respectively.

Simulator test. The simulator test consisted of six repetitions of each of three maneuvers: a level 90° turn to the left or right, a 500 fpm descent, and a 500 fpm descent followed by a power-off stall after level-off. For each maneuver three flight parameters, such as airspeed, were selected for scoring. If the subject exceeded the criterion, performance on that parameter was scored as an error. Thus, the error score on each maneuver could range between 0 and 3. Bank, altitude, and roll-out heading were scored for level turns; altitude, heading, and airspeed for descents; altitude, heading, and airspeed for descents followed by a stall. The following criteria were used for each parameter: bank, $\pm 10^{\circ}$; altitude, ± 100 feet; roll-out heading, $\pm 10^{\circ}$; and airspeed, ± 10 mph. Performance on the last four repetitions of each maneuver was scored independently by the instructor and the experimenter.

Apparatus

Tracking test. The tracking tasks were presented on a 10.2 x 7.6 cm Hewlett-Packard Model 1300A cathode ray tube. A Raytheon 704 computer generated inputs for the tasks, recorded and processed the subject's responses, and timed the trials. The subjects were seated approximately 120 cm from the display.

Two identical Measurement Systems Incorporated Model 435 two-axis spring-centered control sticks were used. Both sticks were modified to permit movement in the left-right dimension only. The control sticks were mounted in adjustable arm rests attached to the subject's chair.

Simulator test. The simulator test was conducted in a Singer-Link General Aviation Trainer (GAT-2). The flight characteristics of the GAT-2 were modified to approximate those of a Piper Cherokee Arrow. All maneuvers were executed in the landing configuration with wheels down and flaps 20° extended. A performance scale, in the form of a booklet with instructions, was used to rate the subjects on selected maneuvers. The scale was similar to that developed by Povenmire, et al. (1970). All training and testing were done without simulator motion.

Procedure

Session 1. Subjects were tested with Form S of the Bennett Test of Mechanical Comprehension (1969). After the subjects finished this test, they began the tracking test. The subjects first completed four single-task trials (commencing the sequence with the nonpreferred hand) followed by 25 dual-task trials. After the subjects finished the dual-task trials, they performed one single-task trial with each hand beginning again with the nonpreferred hand. All trials were 1 min long. A 5-min rest pause was given after Trials 13 and 21; 1-min rest pauses were given after all other trials. The average absolute error scores for each trial were displayed throughout the subsequent rest period. Taped instructions were played to the subjects before the first single- and dual-task trials. The subjects were informed of the number of trials to be completed.

Session 2. Subjects completed audio-visual lessons, loaned by the Jeppesen-Sanderson Company, on flight instruments, basic aerodynamics, and basic flight maneuvers. At the end of this session subjects sat in a front seat of the GAT-2 and listened to a tape describing the controls and instruments. Finally, the subjects were given a workbook and told to complete a written exercise on the flight maneuvers to be practiced in the next lesson.

Session 3. A qualified flight instructor taught subjects pitch control, rudder-aileron coordination, 90° level turns, power management, descents, and stalls, all by instrument reference. During the course of the instruction the subjects attempted six level turns, descents, and stalls. The final four attempts of each maneuver were scored and composed the simulator test.

RESULTS

Performance on the tracking task as a function of practice is shown in Figure 2. The intercorrelations between the 31 tracking trials are given in Table 1.

The tracking data were examined for evidence of the development of timesharing skills using the technique discussed in the Identification of Timesharing Skills Section. The analysis was performed on the data from Trials 4 (single-task), 5 (dual-task), 29 (dual-task), and 30 (single-task). A two-way analysis of variance indicated reliable main effects of trials ($F_{1,56} = 230.91$, $p < .001$) and task load ($F_{1,56} = 771.39$, $p < .001$) and a reliable load by trials interaction ($F_{1,56} = 154.63$, $p < .001$).

The preceding analysis indicates that timesharing skills were learned under dual-task tracking. Because all of the simulator maneuvers presumably require some timesharing, the predictive validities of single-task and dual-task tracking may be compared for evidence that measures of timesharing behavior are better predictors of flight performance than their single-task counterparts.

Several methods of combining simulator data into a unitary performance metric were explored. Principal component and iterated principal factor analyses (Dixon and Brown, 1977) were performed on the maneuver score correlations to identify single factors that could be used as criterion measures. Subsequently, single- and dual-task tracking scores were correlated with the single factor from the principal components analysis, the factors from the iterated principal factor analysis, arbitrary combinations of maneuvers, and the simple sum of the instructors' and observer's

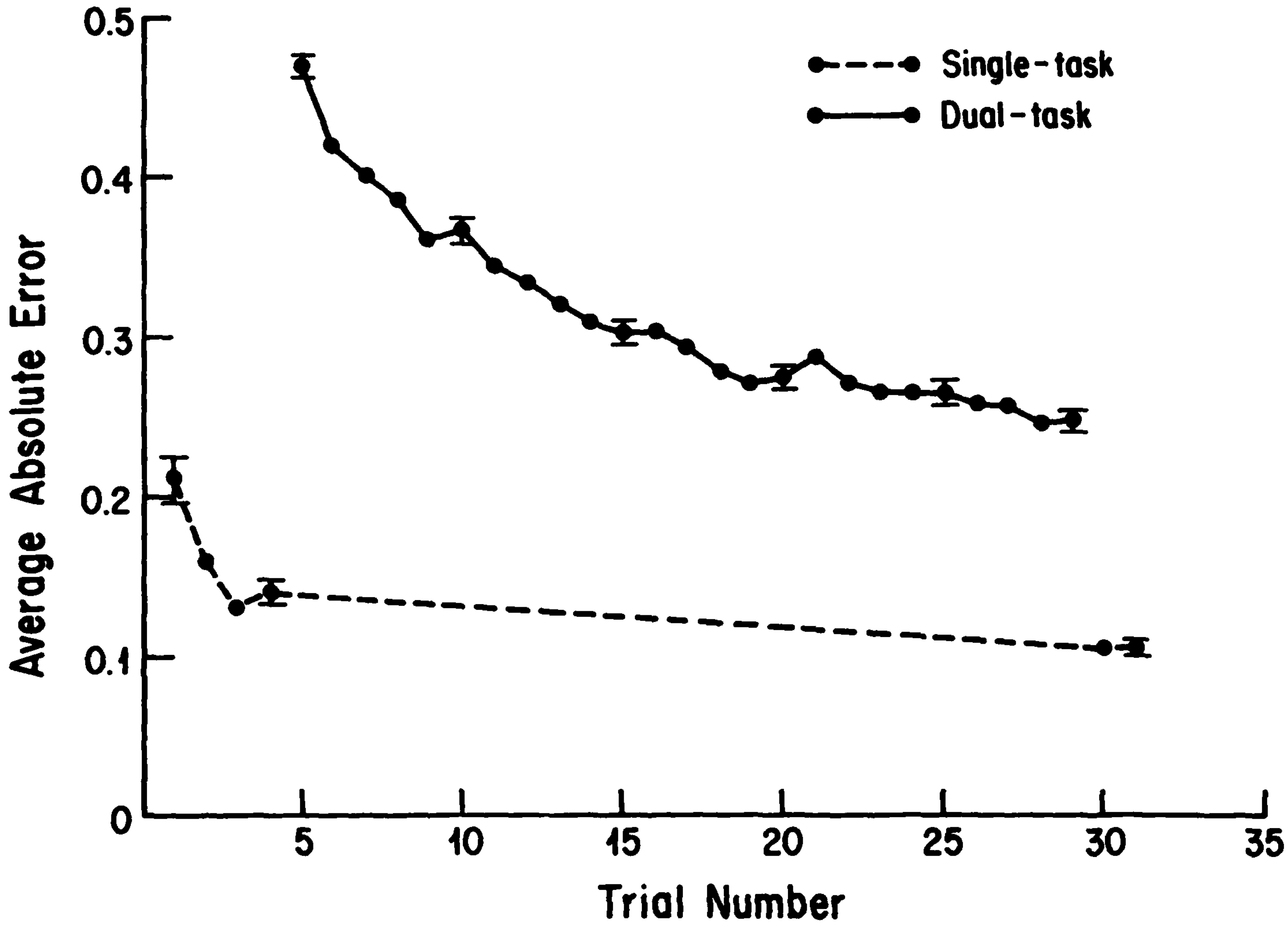


Figure 2. Single- and dual-task tracking performance as a function of practice.

TABLE 1

Intercorrelations of Tracking Trials

Trial Number	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31
1	777	724	242	372	363	355	351	358	352	351	310	200	245	393	291	325	157	322	235	073	353	274	259	353	297	329	302	312	418	330
2		675	287	377	439	404	405	364	401	444	281	226	202	385	386	363	138	339	268	198	335	251	244	307	313	281	275	326	445	425
3			324	263	467	377	372	400	320	372	227	242	307	335	347	358	197	275	241	155	354	189	247	345	287	213	313	288	480	430
4				238	328	232	216	375	487	361	165	257	267	329	260	270	136	168	213	228	306	286	235	364	301	310	175	297	270	232
5					665	671	578	607	557	570	521	488	466	509	453	524	353	332	464	245	432	458	322	452	423	377	381	441	156	288
6						764	557	670	629	664	544	575	520	500	590	515	437	415	453	376	608	516	387	462	447	474	422	434	231	341
7							719	712	599	698	612	617	568	604	631	501	534	531	606	505	605	606	482	493	497	518	373	491	249	344
8								765	594	725	594	549	569	628	578	546	508	543	673	432	524	608	540	505	558	467	529	482	379	441
9									700	760	648	638	662	696	665	651	634	627	740	535	670	667	563	624	665	589	594	623	408	503
10										760	700	720	588	619	610	685	570	594	621	437	647	648	561	598	642	662	604	605	450	442
11											730	775	720	727	853	741	664	660	755	482	710	650	603	660	634	657	602	673	463	463
12												779	771	718	750	725	696	731	740	461	683	626	607	653	670	705	577	592	435	454
13													722	639	780	783	809	729	782	558	775	660	645	683	645	739	678	703	436	441
14														723	806	764	724	737	806	619	784	711	750	781	761	700	645	641	452	461
15															767	766	675	755	743	514	716	711	716	651	714	681	646	762	441	532
16																815	748	791	815	553	755	689	722	721	728	721	659	747	467	523
17																	791	843	790	563	779	739	752	789	830	745	774	799	597	584
18																		792	796	595	747	703	718	712	758	781	685	690	460	418
19																			805	655	804	789	776	783	824	784	770	792	601	600
20																				606	789	790	806	818	795	729	712	752	520	572
21																					632	545	541	588	548	540	530	548	362	366
22																						843	844	778	776	779	757	742	536	524
23																							856	785	856	800	759	786	477	541
24																								798	840	770	784	791	545	549
25																									872	794	785	808	679	632
26																										833	799	815	589	620
27																											780	821	519	516
28																												836	633	683
29																													636	715
30																														782

Note: Decimal points have been omitted.

scores on all maneuvers. In all analyses the pattern of results and the size of the correlations did not vary substantially from those of the simple sum of the maneuver scores. Therefore, all further analyses were conducted only on the sum of the maneuver scores which is referred to as the simulator score. The correlations between single-task and dual-task tracking performance and the simulator score as a function of practice are shown in Figure 3.

Figure 3 indicates a slow improvement in dual-task correlations over the 25 trials while early single-task performance falls monotonically with practice. Because the dual-task correlations fluctuate greatly from trial to trial, it was decided to average dual-task performance over a small number of trials to obtain more stable estimates of the predictive validity of dual-task measures. A close examination of the means and standard deviations of each dual-task trial revealed that the standard deviations within blocks of five trials were very homogeneous; in the worst case the standard deviations varied by a factor of 1.60. Therefore, the scores were averaged over blocks of five trials and new correlations were calculated on the averages. The correlations between the simulator score and dual-task performance on Blocks 1 through 5 were .139, .204, .206, .271, and .287 respectively. The correlations between the simulator score and the first four single-task trials were .194, .103, .059, and -.012 respectively. The correlations with the last two single-task trials (Trials 30 and 31) were .214 and .284.

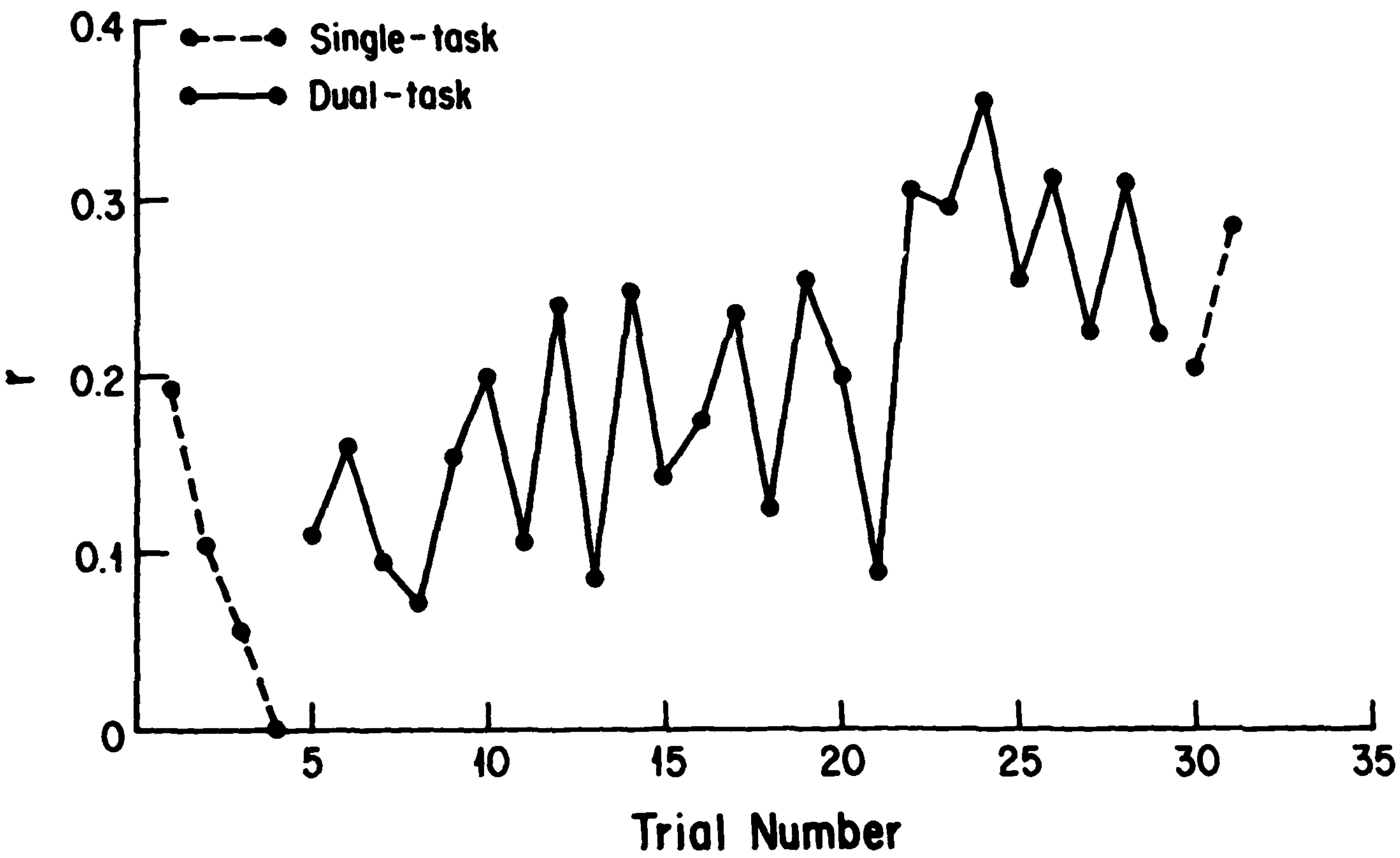


Figure 3. The Pearson product moment correlation between the simulator score and tracking performance as a function of practice.

DISCUSSION

The purpose of this experiment was to compare the predictive validity of single- and dual-task performance measures when timesharing skills were involved in the performance of the dual task. To demonstrate the development of timesharing skills using the technique discussed in the Identification of Timesharing Skills Section, the data must show a reliable main effect of secondary-task load and a reliable secondary-task load by trials interaction. Both these effects were evident in the data. Additionally, single-task performance must remain stable during the period in which timesharing skills are learned. Although the error decreased 4% between Trials 5 and 30 indicating some continued improvement in single-task tracking skills, this change is desirable in that it indicates that performance had not yet reached asymptote and avoided a single-task ceiling effect that would have made interpretation of subsequent improvements in dual-task performance difficult. Additionally, the change in single-task performance was small compared to the corresponding dual-task change during the same period (26%) and it seems evident that timesharing skills developed in this combination.

Although the interaction between secondary-task load and practice is interpreted as evidence for the development of timesharing skills, there are, however, at least two alternative explanations that must be considered. One explanation is that the interaction indicating the development of timesharing skills rests on a number of questionable assumptions about the metric of the dependent variables. Therefore, it may be argued that a transformation of the data could eliminate the interaction. To test this hypothesis, a log transformation, which represents an extreme transformation for these data, was performed on the tracking data. An ANOVA on the transformed data revealed reliable main effects of load ($F_{1,56} = 884.6793$, $p < .01$) and trials ($F_{1,56} = 179.3592$, $p < .01$) and a reliable load by trials interaction ($F_{1,56} = 47.8006$, $p < .01$).

The second explanation is that single-task processing becomes more efficient with practice (consumes less of the operator's attentional resources) even as single-task performance remains unchanged. Norman and Bobrow (1975) have proposed that the performance-resource function--that function which relates performance to the quantity of resources invested--can be differentiated into resource-limited regions in which the quality of performance is proportional to the resources invested and data-limited regions in which performance is unchanged by investment or withdrawal of resources. The explanation of the effects described above would posit that single-task performance is data limited and that the amount of resources required to reach that data limited region becomes progressively less with practice. Thus, the combined resource demands of the two component tasks performed concurrently fall into a resource-limited region and become correspondingly less after practice than before. Therefore, dual-task performance will improve even as data-limited single-task performance remains constant.

However, the processing demands of the tracking task are such that it is unlikely that this task could be described as data limited. The tracking task per se does not impose demands that would exceed any processing characteristics that might represent sources of data limitation (e.g., capacity of short-term memory, speed of response, or resolution of perceptual processing). Therefore, it seems that this explanation cannot account for the data and that the changes in dual-task performance are the result of the development of timesharing skills.

Because the ANOVA shows evidence for the development of timesharing skills under dual-task conditions, the predictive validity of the single-task tracking trials can be compared to that of the dual-task trials. As indicated in the last paragraph of the Results Section, the dual-task tracking scores have no obvious advantage over the late single-task tracking scores (Trials 30 and 31).

However, Figure 3 shows that the correlations between the first four single-task tracking trials and the simulator score decrease monotonically with practice while the dual-task correlations increase with practice. To determine if there is a significant decrease in the predictive validity of the early single-task trials as a function of practice, the correlations were transformed to z scores. A Pearson product-moment correlation then was computed between the z score and the trial number. This correlation was $-.977$ ($p < .05$) indicating that the predictive validity of the single-task tracking correlations decreased significantly with practice. A similar procedure conducted on the 25 dual-task correlations yielded a correlation of .667 ($p < .001$) indicating the predictive validity of dual-task tracking increased significantly with practice. As shown in Figure 3, the predictive validity of the last two single-task trials deviates significantly from the trend established by the preceding single-task trials. They do. however, appear to follow the general trend of the dual-task data and have a magnitude much more similar to that of the dual- rather than the single-task correlations.

To compare the predictive validities of single- and dual-task performance, it is necessary to determine if the scores obtained on Trials 30 and 31 are valid measures of single-task performance after extensive practice or if they have been "contaminated" in some way by the intervening dual-task practice. The pattern of intercorrelations between the single- and dual-task trials given in Table 1 suggests that repeated practice under dual-task conditions probably affected the technique the subjects used to perform the single-task tracking task. Because timesharing skills are learned in the dual-task tracking task, the correlation between the last two single-task trials (Trials 30 and 31) and the last two dual-task trials (Trials 28 and 29) should be approximately equal to or less than the correlation between the last two early single-task trials (Trials 3 and 4) and the first two dual-task trials (Trials 5 and 6). However, the average correlation between Trials 28 and 29 and 30 and 31 is .667 while the average correlation between Trials 3 and 4 and 5 and 6 is only .324. Additionally, the average intercorrelation between Trials 3 and 4 (early single-task) and 30 and 31 (late single-task) is .353. These three intercorrelations indicate that Trials 30 and 31 are less related to the preceding single-task trials and more related to the late dual-task trials than would be anticipated, suggesting that something the subjects used to perform under dual-task conditions may have been employed on the last two single-task trials.

This "something" probably is a dual-task response strategy, which is reflected in the open-loop gain (the ratio of the amplitude of output movement to the input movement) (Wickens and Gopher, 1977). Observation of subjects performing difficult compensatory tracking tasks (such as the one used in this experiment) under single-task conditions indicates that they frequently make large control movements, especially in the early stages of practice. However, under dual-task conditions these large movements often lead to a temporary loss of control on one or both tasks and very large errors. Thus, under dual-task conditions small control movements are best. Under single-task conditions small control movements also lead to small errors and are less fatiguing than large control movements.

It is possible that the subjects initially used large control movements under single-task conditions. The size of the control movements probably was reduced gradually throughout the 25 dual-task trials. When the subjects again performed under single-task conditions (Trials 30 and 31), they employed the same type of strategy as on the immediately preceding trials; they used small control movements, resulting both in a decrease in error and an increase in the predictive validity. Additionally, because strategy is a major determinant of dual-task performance (Damos, 1977), the correlation between late dual-task and late single-task performance easily could have been increased if a dual-task strategy had been employed during single-task performance.

To determine if the subjects did modify their control movements as suggested above, it is necessary to perform a Control Theory Analysis on the tracking data and obtain the open-loop gain on each trial. Such an analysis could not be performed on these data and the explanation given above remains speculative. Additionally, a unambiguous comparison of the predictive validities of dual- versus late single-task performance cannot be made.

It is necessary to discuss the finding that the correlation between averaged dual-task performance and the simulator score increased monotonically with practice. This finding complements previous research (Damos, 1978) showing that the correlation between a dual-task test and the performance on a flight check increased as the subject proceeded through flight training. Damos hypothesized that this unprecedented increase was the result of the development of timesharing skills; as the student progressed through flight training, his timesharing skills were improved. Because the laboratory test measured his timesharing skills, the correlation between this test and the flight check score increased as timesharing skills contributed more to flight performance even thought the inter-test interval increased. If the maneuvers performed in the simulator in the current experiment required timesharing skills and if timesharing skills were learned in dual-task tracking, then the correlation between successive blocks of dual-task tracking and the simulator score should increase, which was observed.

Of course, an alternate explanation is that the inter-test interval decreased with each successive block of dual-task trials. Because the correlation between two measures generally increases as the inter-trial interval decreases, the trend in the dual-task correlation could be attributed solely to a decrease in the inter-test interval. This argument is partially refuted by the trend in the early single-task trials; with each successive trial, the correlation decreased monotonically although the inter-test interval decreased.

In summary, the data reported in this experiment show the development of timesharing skills under dual-task conditions. The correlation between successive blocks of dual-task trials and performance in a simulator increased while the correlation between successive early single-task trials and simulator performance decreased. However, the last two single-task trials showed a large increase in predictive validity, approximating that of the final block of dual-task trials. Although this increase may indicate that dual-task measures do not correlate more highly with simulator performance than single-task measures, the pattern of intercorrelations between single- and dual-task tracking scores suggest that a multiple-task performance strategy may have been employed on the last two single-task trials resulting in the large increase in correlation.

REFERENCES

Alvares, K.M. and Hulin, C.L. Two explanations of temporal changes in ability-skill relationships: A literature review and theoretical analysis. Human Factors, 1972, 14, 295-308.

Bennett, G.K. Bennett Mechanical Comprehension Test. Form S. New York: The Psychological Corporation, 1969.

Damos, D.L. Development and transfer of timesharing skills. Savoy, Ill.: University of Illinois at Urbana-Champaign, Aviation Research Laboratory, TR ARL-77-11/AFOSR-77-10, 1977.

Damos, D.L. Residual attention as a predictor of pilot performance. Human Factors, 1978, 20, 435-440.

Dixon, W.J. and Brown, M.B. (Eds.) BMDP biomedical computer programs P-series. Berkeley: University of California Press, 1977.

Jacobs, R.S. Simulator cockpit motion and the transfer of initial flight training. Savoy, Ill.: University of Illinois at Urbana-Champaign, Institute of Aviation, Aviation Research Laboratory, TR ARL-76-8/AFOSR-76-4, 1976.

Jennings, A.E. and Childs, W.D. An investigation of time-sharing ability as a factor in complex performance. Human Factors, 1977, 19, 535-548.

Lintern, G. Transfer of landing skill after training with supplementary visual cues. Savoy, Ill.: University of Illinois at Urbana-Champaign, Department of Psychology, Technical Report Eng Psy-78-3/AFOSR-78-2, 1978.

Melton, A.W. (Ed.) Apparatus tests. Washington, D.C.: Government Printing Office, Army Air Force Aviation Psychology Program Research Report No. 4, 1947.

Norman, D.A. and Bobrow, D.G. On data-limited and resource-limited processes. Cognitive Psychology, 1975, 7, 44-64.

North, R.A. and Gopher, D. Measures of attention as predictors of flight performance. Human Factors, 1976, 18, 1-14.

Povenmire, H.K., Alvares, K.M., and Damos, D.L. Observer-observer flight check reliability. Savoy, Ill.: University of Illinois at Urbana-Champaign, Institute of Aviation, Aviation Research Laboratory, TR LF-70-2, 1970.

Sverko, B. Individual differences in time-sharing performance. Savoy, Ill.: University of Illinois at Urbana-Champaign, Institute of Aviation, Aviation Research Laboratory, TR ARL-77-4/AFOSR-77-4, 1977.

Trankell, A. The psychologist as an instrument of prediction. Journal of Applied Psychology, 1959, 43, 170-175.

Wickens, C.D. and Gopher, D. Control Theory measures of tracking as indices of attention allocation strategies. Human Factors, 1977, 19, 349-366.